Acta Math. Hung.
49 (1—2) (1987), 251—259.

ON LOCALLY REPEATED VALUES OF CERTAIN ARITHMETIC FUNCTIONS. II

P. ERDŐS (Budapest), member of the Academy,
C. POMERANCE* (Morristown) and A. SÁRKÖZY (Budapest)

§1. Introduction

Let $\nu(n)$ denote the number of distinct prime factors of n. In Part I of this paper [5], we studied the equation

(1.1) $$n+\nu(n) = m+\nu(m), \quad n \neq m$$

and generalizations where ν is replaced with a more general arithmetic function. In this part, we study the distribution of the n for which $\nu(n) \approx \nu(n+1)$ and the n for which $\Phi(n)=\Phi(n+1)$, where Φ is Euler's function.

It seems reasonable to conjecture that there is a positive constant c, such that the number of $n \leqq x$ with

(1.2) $$\nu(n) = \nu(n+1)$$

is $(c_1+o(1))x/\sqrt{\log\log x}$. Indeed, from the Erdős—Kac theorem, most integers n satisfy

(1.3) $$|\nu(n) - \log\log n| \leqq K\sqrt{\log\log n}$$

where K is some large constant. In fact, the asymptotic density of the n which satisfy (1.3) is exactly

$$\frac{1}{\sqrt{2\pi}} \int_{-K}^{K} e^{-t^2/2}\, dt,$$

which is nearly 1 if K is large. Thus if n and $n+1$ both satisfy (1.3) and if we view $\nu(n)$ and $\nu(n+1)$ as "independent events", then the "probability" that (1.2) holds should be at least $(2K\sqrt{\log\log n})^{-1}$. Summing these probabilities would then give order of magnitude $x/\sqrt{\log\log x}$ solutions n of (1.2) with $n \leqq x$, thus supporting the conjecture. A refinement of this heuristic argument even suggests that $c_1=(2\sqrt{\pi})^{-1}$.

It is not even known, however, if (1.2) has infinitely many solutions. Our principal result in this paper is that a slight weakening of (1.2) has at least the "correct" order of magnitude for the number of solutions $n \leqq x$.

THEOREM 1. *There are absolute constants $c_2, c_3 > 0$ such that for $x \geqq 3$, the number of $n \leqq x$ with*

(1.4) $$|\nu(n) - \nu(n+1)| \leqq c_2$$

is at least $c_3 x/\sqrt{\log\log x}$.

---

* Partially supported by an NSF grant. On leave from the University of Georgia.

The proof of Theorem 1 is based on the fundamental lemma of the combinatorial sieve. By using Selberg's lower bound sieve (see Halberstam and Richert [6], Theorem 7.4 which gives Lemma 1 below with $u=4.43$) and a more careful argument, Theorem 1 can be proved with $c_2=3$. However, the proof would be longer and not involve any essentially new ideas, so we present here only the simpler version in Theorem 1.

It is to be remarked that our proof of Theorem 1 easily gives the same result with Ω in place of ν, where $\Omega(n)$ is the number of prime factors of n counting multiplicity. If $d(n)$ denotes the number of divisors of n, our proof also shows that the number of $n\leqq x$ with $d(n)/d(n+1)=2^i$ where i is an integer with $|i|\leqq c_2$ is at least $c_3x/\sqrt{\log\log x}$. Recently, Heath-Brown [7] proved that $d(n)=d(n+1)$ has at least $c_4x/(\log x)^7$ solutions $n\leqq x$ where $c_4>0$. He announced that his method also works for $\Omega(n)=\Omega(n+1)$. It is not unlikely that some refinement of his idea would also work for (1.2).

A result somewhat weaker than Theorem 1 can be immediately obtained using a special case of a theorem of Barban and Vinogradov (see Elliott [2], Theorem 20.1). From this theorem, for $x>10^7$ and $z>0$,

$$\frac{1}{x}\cdot \#\{n\leqq x\colon |\nu(n+1)-\nu(n)|\leqq z\sqrt{2\log\log x}\}=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-z}^{z} e^{-t^2/2}\,dt+O\left(\frac{1}{\sqrt{\log\log x}}\left(1+\frac{\log\log\log x}{e^{t^2/2}\log\log\log\log x}\right)\right).$$

Applying this result to values of z near 0 gives

$$\#\left\{n\leqq x\colon |\nu(n+1)-\nu(n)|\leqq c_5\frac{\log\log\log x}{\log\log\log\log x}\right\}\asymp\frac{x\log\log\log x}{\sqrt{\log\log x}\,\log\log\log\log x},$$

for some constant $c_5>0$. We are endebted to R. R. Hall and G. Tenenbaum for this observation.

Some changes in the proof of Theorem 1 give the following result which we state without proof.

THEOREM 1′. *For each positive integer k, there are absolute constants $c_2(k)$, $c_3(k)>0$ such that for $x\geqq 3$, the number of $n\leqq x$ with*

$$\max\{\nu(n),\nu(n+1),\dots,\nu(n+k)\}-\min\{\nu(n),\nu(n+1),\dots,\nu(n+k)\}\leqq c_2(k)$$

is at least $c_3(k)x/(\operatorname{loglog} x)^{k/2}$.

In the third paper in this series, we shall complement Theorem 1 with an upper bound result of the same order of magnitude and also give an upper bound for the frequency of solutions of (1.1).

We shall complete this paper with a result concerning Euler's function $\Phi(n)$. In [3], it is shown that the asymptotic density of the n with $\Phi(n)<\Phi(n+1)$ is 1/2 and the same for the n with $\Phi(n)>\Phi(n+1)$. Thus as a corollary, the number of $n\leqq x$ with

$$\Phi(n)=\Phi(n+1) \tag{1.5}$$

is $o(x)$. The following result implies, for example, the stronger assertion that the solutions of (1.5) have a bounded sum of reciprocals.

THEOREM 2. *For large x, the number of solutions of* (1.5) *not exceeding x is at most* $x/\exp\{(\log x)^{1/3}\}$.

The proof of Theorem 2, which is based on the argument in [8], can also be used to show the same result for the equation $\sigma(n)=\sigma(n+1)$, where σ is the sum of the divisors function. We conjecture that for every $\varepsilon>0$ and $x\geqq x_0(\varepsilon)$ the equations $\Phi(n)=\Phi(n+1)$, $\sigma(n)=\sigma(n+1)$ each have at least $x^{1-\varepsilon}$ solutions $n\leqq x$. We cannot prove, however, that there are even infinitely many solutions for either equation.

§2. Preliminaries for Theorem 1

In this section, we record three results which will be used in the proof of Theorem 1.

LEMMA 1. *Let $p_0(n)$ denote the least odd prime factor of n if n is not a power of 2 and let $p_0(2^k)=1$. Then there are real numbers $u>1$, x_0, and $c_6>0$ such that if a, b, a_0, b_0 are nonnegative integers satisfying*

$$ab\neq 0,\quad ab_0-a_0b=1$$

and if $x\geqq \max(x_0, a^u, b^u)$, *then*

$$\#\{n\leqq x\colon\ p_0\big((an+a_0)(bn+b_0)\big)>x^{1/u},\ (an+a_0)(bn+b_0)\not\equiv ab \bmod 2\}>$$

$$> c_6\frac{ab}{\Phi(a)\Phi(b)}\frac{x}{\log^2 x}.$$

PROOF. This follows from the "fundamental lemma" of the combinatorial sieve (see Halberstam and Richert [6], Theorem 2.5). Note that if $2|ab$, then the sieve result implies that we may insist that $(an+a_0)(bn+b_0)$ be odd, while if $2\nmid ab$, then $(an+a_0)(bn+b_0)$ is even for all n.

LEMMA 2. *Let $\pi(x,t)$ denote the number of $n\leqq x$ with $\nu(n)=t$. Then uniformly for $x\geqq 3$ and integers t satisfying*

$$|t-\log\log x|<(\log\log x)^{2/3}, \tag{2.1}$$

we have

$$\pi(x,t)=\frac{e^{-c^2/2}}{\sqrt{2\pi}}\frac{x}{\sqrt{\log\log x}}\left(1+O\left(\frac{|c+c^3|}{\sqrt{\log\log x}}\right)+O\big((\log\log x)^{-1}\big)\right)$$

where c is defined by the equation: $t=\log\log x+c\sqrt{\log\log x}$.

PROOF. From a result of Sathe [9] and Selberg [10], for any $B>0$ we have uniformly for $x\geqq 3$ and integers t with $1\leqq t\leqq B\log\log x$,

$$\pi(x,t)=\frac{x}{\log x}\frac{(\log\log x)^{t-1}}{(t-1)!}F\left(\frac{t-1}{\log\log x}\right)\left(1+O\left(\frac{t}{(\log\log x)^2}\right)\right)$$

where

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

If t satisfies (2.1), then

$$F\left(\frac{t-1}{\log\log x}\right) = 1+O\left(\frac{|c|}{\sqrt{\log\log x}}\right)+O((\log\log x)^{-1}),$$

$$\frac{(\log\log x)^{t-1}}{(t-1)!} = \frac{e^{-c^2/2}}{\sqrt{2\pi}}\frac{\log x}{\sqrt{\log\log x}}\left(1+O\left(\frac{|c+c^3|}{\sqrt{\log\log x}}\right)+O\big((\log\log x)^{-1}\big)\right),$$

which proves the lemma. Note that a somewhat weaker version of this lemma follows from [4].

LEMMA 3. *Let us put*

$$S(n) = \sum_{p|n} \frac{1}{p}.$$

There are numbers x_1 and Δ, $\eta>0$ such that if $x>x_1$, then there are at least ηx nitegers $n\leqq x$ with

(2.2) $$S(n) \leqq \frac{1}{10} \quad \text{and} \quad |\nu(n)-\log\log x| \leqq \Delta\sqrt{\log\log x}\,.$$

PROOF. For any t, let $D(t)$ denote the asymptotic density of the integers n with $S(n)\leqq t$. By the Erdős—Wintner theorem (see Elliott [2], Theorem 5.1) $D(t)$ exists for every t. It is easy to see that $D(t)$ is strictly increasing on $[0, \infty)$. Thus $D(1/10)>0$.

By the Erdős—Kac theorem (see Elliott [2], Theorem 12.3) the number of $n\leqq x$ with $|\nu(n)-\log\log x|>\Delta\sqrt{\log\log x}$ is

$$\left(\frac{2}{\sqrt{2\pi}}\int_\Delta^x e^{-t^2/2}\,dt+o(1)\right)x.$$

Let $0<\eta<D(1/10)$ be arbitrary and choose Δ so that

$$\frac{2}{\sqrt{2\pi}}\int_\Delta^x e^{-t^2/2}\,dt < D\left(\frac{1}{10}\right)-\eta.$$

Then for large x, at least ηx integers $n\leqq x$ satisfy (2.2).

§3. Proof of Theorem 1

Basically, the idea of the proof of Theorem 1 is to construct many pairs $n, n+1$ from a fixed choice of integers a, b where $b|n$, $a|n+1$, $\nu(a)=\nu(b)$, and $\nu(n/b)$, $\nu((n+1)/a)$ are small. Then $\nu(n)\approx\nu(n+1)$. To show there are many such n for a given choice of a, b, we use Lemma 1. To show there are many pairs a, b, we use Lemmas 2 and 3.

Let x be large. Let $u>1$ denote the number defined in Lemma 1 and let $v=u+2$. Let $\mathcal{A}$ denote the set of integers a satisfying

$$a \leqq x^{1/v}, \quad S(a) \leqq \frac{1}{10}, \quad |\nu(a)-\log\log x| \leqq 2\Delta\sqrt{\log\log x}\,, \tag{3.1}$$

where S and Δ are defined in Lemma 3. Then (with η given in Lemma 3)

$$\sum_{a\in A}\frac{1}{a} > \frac{\eta}{2v}\log x. \tag{3.2}$$

Indeed, if $x^{1/4v}<y\leqq x^{1/v}$ and x is large, then by Lemma 3, the number of members of $\mathcal{A}$ that do not exceed y is at least ηy. (We use the fact that $\log\log y=\log\log x+ +O(1)$). Thus (3.2) follows from partial summation.

For $a\in\mathcal{A}$, let $\mathcal{B}(a)$ denote the set of integers b with

$$b \leqq x^{1/v}, \quad (a, b) = 1, \quad \nu(b) = \nu(a). \tag{3.3}$$

Finally, for $a\in\mathcal{A}$ and $b\in\mathcal{B}(a)$, let $g(a, b)$ denote the number of integer solutions y_1, y_2 of

$$ay_1-by_2 = 1, \; 0 < y_1 \leqq \frac{x}{a}, \; p_0(y_1y_2) > x^{1/v}, \; ab \not\equiv y_1y_2 \mod 2. \tag{3.4}$$

For such a quadruple a, b, y_1, y_2 let

$$n = by_2, \quad n+1 = ay_1. \tag{3.5}$$

Then $n+1\leqq x$ and by (3.1), (3.3), and (3.4), we have $(a, y_1)=(b, y_2)=1$. Thus from (3.3), (3.4), and (3.5)

$$|\nu(n+1)-\nu(n)| = |\nu(a)+\nu(y_1)-\nu(b)-\nu(y_2)| = |\nu(y_1)-\nu(y_2)| < v+1.$$

We shall then take $c_2=v+1$ in Theorem 1.

We next note that for a given integer n, there is at most one quadruple a, b, y_1, y_2 satisfying (3.1), (3.3), (3.4), and (3.5). Thus if $f(x)$ denotes the number of $n\leqq x$ satisfying (1.4) with $c_2=v+1$, then

$$f(x) \geqq \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}(a)} g(a, b). \tag{3.6}$$

Note that Lemma 1 immediately gives a lower bound for each $g(a, b)$. Indeed, if positive integers y_1, y_2 satisfy $ay_1-by_2=1$, then y_1, y_2 are of the form

$$y_1 = bm+b_0, \quad y_2 = am+a_0,$$

where $m=0, 1, 2, \ldots$,

$$0 \leqq b_0 < b, \quad 0 \leqq a_0 < a, \quad \text{and} \quad ab_0 - ba_0 = 1.$$

Thus from Lemma 1,

$$g(a, b) > c_6 \frac{ab}{\Phi(a)\Phi(b)} \frac{x/ab}{\log^2(x/ab)} - 1 \geqq c_6 \frac{x}{ab \log^2 x}.$$

Therefore, from (3.6)

$$f(x) \geqq c_6 \frac{x}{\log^2 x} \sum_{a \in \mathcal{A}} \frac{1}{a} \sum_{b \in \mathcal{B}(a)} \frac{1}{b}. \tag{3.7}$$

We now estimate the inner sum. If $\mathcal{B}(a)(y)$ denotes the number of members of $\mathcal{B}(a)$ below y, then, for $y \leqq x^{1/\nu}$,

$$\mathcal{B}(a)(y) = \sum_{\substack{b \leqq y \\ (a,b)=1 \\ \nu(b)=\nu(a)}} 1 \geqq \sum_{\substack{b \leqq y \\ \nu(b)=\nu(a)}} 1 - \sum_{p|a} \sum_{\substack{b \leqq y \\ p|b \\ \nu(b)=\nu(a)}} 1 \geqq$$

$$\geqq \pi(y, \nu(a)) - \sum_{p|a} \left(\pi\left(\frac{y}{p}, \nu(a)\right) + \pi\left(\frac{y}{p}, \nu(a)-1\right) \right).$$

For $x^{1/4\nu} < y \leqq x^{1/\nu}$ and $p \leqq \sqrt{y}$,

$$\pi\left(\frac{y}{p}, \nu(a)\right) \sim \pi\left(\frac{y}{p}, \nu(a)-1\right) \sim \frac{1}{p} \pi(y, \nu(a))$$

uniformly, by Lemma 2 and (3.1). Therefore for large x, and $x^{1/4\nu} < y \leqq x^{1/\nu}$,

$$\mathcal{B}(a)(y) \geqq \left(1 - \sum_{p|a} \frac{3}{p}\right) \pi(y, \nu(a)) - 2 \sum_{\substack{p|a \\ p > \sqrt{y}}} \frac{y}{p} = (1 - 3S(a))\pi(y, \nu(a)) + O(\sqrt{y}).$$

Thus from (3.1), we have

$$\mathcal{B}(a)(y) \geqq \frac{1}{2} \pi(y, \nu(a))$$

for large x and $x^{1/4\nu} < y \leqq x^{1/\nu}$. From this inequality, (3.1), and Lemma 2, we have

$$\sum_{b \in \mathcal{B}(a)} \frac{1}{b} > \left(4\sqrt{2\pi}\, \nu e^{2A^2}\right)^{-1} \frac{\log x}{\sqrt{\log \log x}}$$

for large x. Combined with (3.2), and (3.7), we thus have

$$f(x) \gg \frac{x}{\sqrt{\log \log x}},$$

which was to be proved.

§4. Proof of Theorem 2

In this section, we outline the proof of Theorem 2. It is nearly the same as the proof in [8] dealing with amicable numbers.

For simplicity of notation, we put

$$l = \exp\{(\log x)^{1/3}\},\ L = \exp\left\{\frac{1}{8}(\log x)^{2/3}\log\log x\right\}.$$

Let $P(n)$ denote the largest prime factor of n. From deBruijn's estimate [1], the number of $n \leqq x$ with $P(n) < L^2$ is $o(x/l)$. Thus we may assume that

(i) $P(n) \geqq L^2$ and $P(n+1) \geqq L^2$.

It is easy to see that the number of $n \leqq x$ divisible by a non-trivial power exceeding l^3 is $o(x/l)$, so we may assume that

(ii) if k^a divides n or $n+1$ where $a \geqq 2$, then $k^a \leqq l^3$.

Now we show that we may assume that

(iii) $n/P(n) \geqq L$, $(n+1)/P(n+1) \geqq L$.

Let $n=mp$, $n+1=m'p'$ where $p=P(n)$, $p'=P(n+1)$. From (i), (ii), and (1.5) we have

$$\Phi(m)(p-1) = \Phi(m')(p'-1).$$

Using this equation and $mp+1=m'p'$, we have

$$p'\big(\Phi(m)m' - m\Phi(m')\big) = \Phi(m) - m\Phi(m') + m\Phi(m). \tag{4.1}$$

Assuming $n \geqq 4$, (1.5) implies that $m>1$ and $m'>1$. Thus the parenthetical expression in (4.1) is not 0, for otherwise

$$\frac{\Phi(m)}{m} = \frac{\Phi(m')}{m'},$$

contradicting $(m, m')=1$, $m>1$, $m'>1$. Thus m, m' determine p' and p. So the number of $n \leqq x$ for which (1.5), (i), and (ii) hold, but (iii) fails is at most the number of pairs m, m' where either

$$m < L,\ m' \leqq (x+1)/L^2 \quad\text{or}\quad m \leqq x/L^2,\ m' < L,$$

which is $O(x/L) = o(x/l)$.

Continuing with the notation $m=n/P(n)$, $m'=(n+1)/P(n+1)$, we now show we may assume that

(iv) $P\big(\Phi(m)\big) \geqq l^4$, $P\big(\Phi(m')\big) \geqq l^4$.

Suppose $P\big(\Phi(m)\big) < l^4$. We have

$$\Phi(m) = \prod_{q^a \| m} (q-1)q^{a-1} = a_1 a_2 \dots a_t, \tag{4.2}$$

where q denotes prime and each a_i is some $q-1$ or q. Thus the function Φ not only maps m to the integer $\Phi(m)$ but also gives a factorization of $\Phi(m)$ as $a_1 a_2 \dots a_t$

(order of factors is unimportant) where at most one $a_i=1$. It is easy to see (cf. [8]) that there are never 3 distinct integers m_1, m_2, m_3 such that not only are $\Phi(m_1)=\Phi(m_2)=\Phi(m_3)$, but the factorizations $a_1a_2\dots a_t$ given by (4.2) are the same.

If $f(k)$ denotes the number of unordered factorizations of k into factors exceeding 1, then the number of unordered factorizations of k into factors where at most one factor is 1 is $2f(k)$ (for $k>1$). Let $N(z)$ denote the number of m with $1<\Phi(m)\le z$ and $P(\Phi(m))<l^4$. Thus

$$N(z)\le 4\sum_{\substack{k\le z\\ P(k)<l^4}} f(k).$$

This sum is identical with the sum in the first display on p. 186 in [8]. From the argument there (see (6)), for all large x and $z\ge L$,

$$N(z)\le z/l^2. \tag{4.3}$$

If $n\le x$ satisfies (1.5) and (i)—(iii), then

$$L\le m=\frac{n}{P(n)}\le\frac{x}{P(n)}.$$

Thus the number of such n with the first inequality in (iv) failing is by (4.3) at most

$$\sum_{p\le x/L} N(x/p)\le\sum_{p\le x/L}\frac{x}{pl^2}=o(x/l).$$

Similarly, the number of n which the second inequality in (iv) fails is $o(x/l)$. Thus we may assume (iv).

Finally, we may assume that

(v) $P(n)>P(n+1)$.

Indeed, the case $P(n)<P(n+1)$ can be treated in the same way so that there is no loss of generality in assuming (v).

Let $n=mp$, $n+1=m'p'$ where $p=P(n)$, $p'=P(n+1)$ and assume $n\le x$, n satisfies (1.5) and (i)—(v). From (iv) there is a prime $r\ge l^4$ with $r|\Phi(m)$, so that $r|\Phi(n)=\Phi(n+1)$. Thus from (ii) there are primes q, q' with $q|m, q'|n+1$ with $q\equiv q'\equiv 1 \bmod r$. Thus

$$n+1=mp+1\equiv 0 \bmod q'. \tag{4.4}$$

Since $m|n$, $q'|n+1$, we have $(m,q')=1$, so that (4.4) puts p in a certain residue class $a(m,q') \bmod q'$. Also, $p>q'$ by (v). Thus the number of $n\le x$ which satisfy (1.5) and (i)—(v) is at most

$$\sum_{r\ge l^4}\sum_{\substack{q\equiv 1(r)\\ q\le x}}\sum_{\substack{m\equiv 0(q)\\ m\le x}}\sum_{\substack{q'\equiv 1(r)\\ q'\le x+1}}\sum_{\substack{p\equiv a(m,q')(q')\\ q'<p\le x/m}} 1\le$$

$$\le\sum_r\sum_q\sum_m\sum_{q'}\frac{x}{mq'}\ll\sum_r\sum_q\sum_m\frac{x\log x}{rm}\ll$$

$$\ll\sum_r\sum_q\frac{x\log^2 x}{rq}\ll\sum_r\frac{x\log^3 x}{r^2}\ll\frac{x\log^3 x}{l^4}=o(x/l).$$

This completes the proof of Theorem 2.

References

[1] N. G. de Bruijn, On the number of integers $\leqq x$ and free of prime factors $>y$, *Nederl. Akad. Wetensch. Proc. Ser. A*, **54** (1951), 50—60.

[2] P. D. T. A. Elliott, *Probabilistic Number Theory*, vols. I and II, Springer-Verlag (New York, 1979, 1980).

[3] P. Erdős, On a problem of Chowla and some related problems, *Proc. Cambridge Philos. Soc.*, **32** (1936), 530—540.

[4] P. Erdős, On the integers having exactly K prime factors, *Ann. Math.*, **49** (1948), 53—66.

[5] P. Erdős, C. Pomerance, and A. Sárközy, On locally repeated values of certain arithmetic functions, I, *J. Number Theory* (to appear).

[6] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press (London, 1974).

[7] D. R. Heath-Brown, The divisor function at consecutive integers (to appear).

[8] C. Pomerance, On the distribution of amicable numbers, II, *J. reine angew. Math.*, **325** (1981), 183—188.

[9] L. G. Sathe, On a problem of Hardy on the distribution of integers having a given number of prime factors, *J. Indian Math. Soc. (N. S.)*, **17** (1953), 63—141; **18** (1954), 27—81.

[10] A. Selberg, Note on a paper by L. G. Sathe, *J. Indian Math. Soc. (N. S.)*, **18** (1954), 83—87.

(Received February 27, 1985)

P. ERDŐS AND A. SÁRKÖZY
MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, REÁLTANODA U. 13—15.
H—1053

C. POMMERANCE
BELL COMMUNICATIONS RESEARCH
MORRISTOWN, NJ 07960
USA

86-2052 — Szegedi Nyomda — F. v.: Surányi Tibor igazgató