COLLOQUIUM MATHEMATICUM

VOL. XVII 1967 FASC. 2

SOME REMARKS ON THE ITERATES OF THE φ AND σ FUNCTIONS

BY

P. ERDÖS (BUDAPEST)

Put $\sigma_1(n) = \sigma(n)$, $\varphi_1(n) = \varphi(n)$ and, for $k > 1$, $\sigma_k(n) = \sigma_1(\sigma_{k-1}(n))$, $\varphi_k(n) = \varphi_1(\varphi_{k-1}(n))$.

Schinzel conjectured that for every k

$$\liminf \frac{\sigma_k(n)}{n} < \infty. \tag{1}$$

Mąkowski and Schinzel [2] proved (1) for $k = 2$. In fact, they showed (among others) that

$$\liminf \frac{\sigma_2(n)}{n} = 1 \quad \text{and} \quad \limsup \frac{\varphi_2(n)}{n} = \frac{1}{2}.$$

At present, I cannot prove (1) for $k = 3$, but I show the following differences between the cases $k = 2$ and $k = 3$. Denote by $N_\varphi(k, a, x)$ the number of integers $n \leqslant x$ for which

$$\varphi_k(n) > an,$$

and by $N_\sigma(k, a, x)$ the number of integers $n \leqslant x$ for which

$$\sigma_k(n) < an.$$

THEOREM 1. *For every $a < \frac{1}{2}$, arbitrarily small $\varepsilon > 0$ and arbitrarily large t we have for $x > x_0(a, t, \varepsilon)$ the inequalities*

$$\frac{x}{\log x}(\log\log x)^t < N_\varphi(2, a, x) < \frac{x}{\log x}(\log x)^\varepsilon; \tag{2}$$

further, for every $a > 0$ and $\varepsilon > 0$, we have for $x > x_0(a, \varepsilon)$

$$N_\varphi(3, a, x) < \frac{x}{(\log x)^2}(\log x)^\varepsilon. \tag{3}$$

THEOREM 2. *We have for every t if $x > x_0(t)$*

$$N_\sigma(2, 2, x) > \frac{x}{\log x} (\log\log x)^t \tag{4}$$

and for every $a > 0$ and $\varepsilon > 0$ if $x > x_0(\varepsilon, a)$

$$N_\sigma(2, a, x) < \frac{x}{\log x} (\log x)^\varepsilon, \qquad N_\sigma(3, a, x) < \frac{x}{(\log x)^2} (\log x)^\varepsilon . \tag{5}$$

For $n > 2$ we have $\varphi_2(n) < n/2$, thus, in Theorem 1, $a < \frac{1}{2}$ is the best possible.

Before I prove these theorems, I would like to make a few remarks. Let $p > 2$ be any prime (throughout this paper p, q and r will denote primes). Denote by Q_1 the set of all primes $q_1^{(1)} < q_2^{(1)} < \dots$ satisfying $q_i^{(1)} \equiv 1 \pmod p$. Denote by Q_2 the set of primes $q_1^{(2)} < q_2^{(2)} < \dots$ for which $q_i^{(2)} \equiv 1 \pmod{q_j^{(1)}}$ for at least one j but which are not in Q_1. Generally, Q_k denotes the set of primes $q_1^{(k)} < q_2^{(k)} < \dots$ for which $q_i^{(k)} \equiv 1 \pmod{q_j^{(k-1)}}$ for at least one j but which do not belong to $\bigcup_{l=1}^{k-1} Q_l$; in other words, $q_i^{(k)} \not\equiv 1 \pmod{q_j^{(l)}}$ for every j and $l < k-1$. Put

$$Q^{(k)} = \bigcup_{l=1}^{k} Q_l, \qquad Q_\infty = \bigcup_{l=1}^{\infty} Q_l;$$

$\bar{Q}^{(k)}$ and $\bar{Q}_\infty$ denote the sets of primes which do not belong to $Q^{(k)}$ and Q_∞ respectively. $N_x(Q)$ denotes the number of elements not exceeding x of the set Q. It follows from the prime number theorem for arithmetic progressions that

$$N_x(Q_1) = \big(1 + o(1)\big) \frac{x}{(p-1)\log x} .$$

It easily follows from the prime number theorem for arithmetic progressions and the sieve of Eratosthenes that

$$N_x(Q_2) = \big(1 + o(1)\big) \frac{x}{\log x} .$$

By using Brun's method we easily obtain the following stronger result ($c_1, c_2, \dots$ are positive absolute constants):

$$N_x(Q^{(2)}) < c_1 x/(\log x)^{1+1/(p-1)} . \tag{6}$$

The proof of (6) is quite straightforward and can be left to the reader. I have not proved that $N_x(\bar{Q}^{(2)})$ tends to infinity as $x \to \infty$, but this should perhaps be possible by Linnik's method [1]. In other words, the problem (**P 595**) is to prove that there are infinitely many primes r for which

$$r \not\equiv 1 (\operatorname{mod} p) \quad \text{and} \quad r \not\equiv 1 (\operatorname{mod} q_i^{(1)}), \quad i = 1, 2, \ldots$$

It is easy to deduce from (6) by using Brun's method that

$$N_x(\bar{Q}^{(3)}) < c_2 x/(\log x)^2. \tag{7}$$

Very likely there are infinitely many primes in each Q_k and also in $\bar{Q}_\infty$. The problem of the existence of infinitely many primes in $\bar{Q}_\infty$ and Q_k is connected with the following question. Let $p_1^{(1)} = 2 < p_2^{(1)} < \ldots < p_r^{(1)}$ be a finite set of primes. We define inductively a set of primes as follows. By $p_1^{(2)} < p_2^{(2)} < \ldots$ we denote the set of primes, for which $p_i^{(2)} - 1$ is composed entirely of the $p_i^{(1)}$'s. Generally, the $p_i^{(k)}$ are the primes for which $p_i^{(k)} - 1$ is composed entirely of the $p_i^{(l)}$, $l < k$. It seems likely that for every k there are primes $p_i^{(k)}$ (perhaps infinitely many), but nothing is known about this. It is not difficult to deduce from (7) that the number of the $p_i^{(k)}$, $i = 1, 2, \ldots, k = 1, 2, \ldots,$ not exceeding x is less than $c_3 x/(\log x)^2$ but very likely this is a very poor upper bound.

We can prove that for every $\varepsilon > 0$ for all but $\sigma(x)$ integers $n < x$

$$\sigma_k(n) \equiv 0 \Big(\operatorname{mod} \prod_{p < (\log\log x)^{k-\varepsilon}} p\Big).$$

The same result holds for $\varphi_k(n)$. Further we can show that if we neglect a sequence of density 0, then

$$\frac{\sigma_k(n)}{\sigma_{k-1}(n)} = \big(1 + o(1)\big) \frac{\varphi_{k-1}(n)}{\varphi_k(n)} = \big(1 + o(1)\big) k e^{\gamma} \log\log\log n$$

but we do not prove these results in this note.

We will only prove Theorem 1 since the proof of Theorem 2 is similar, but even in the proof of Theorem 1 we will not always give all the details. First we discuss to what extent our theorems are the best possible. We have, for $n > 2$, $\varphi_2(n) < n/2$; thus in Theorem 1 the number $\frac{1}{2}$ cannot be replaced by any greater number. It seems very hard to give an asymptotic formula for $N_\varphi(2, a, x)$ or $N_\sigma(2, a, x)$ (see (3)) and the second inequality of (5) can perhaps be improved (**P 596**).

Now we discuss (4). It is best possible in the sense that $a = 2$ cannot be replaced by any smaller number. We outline the proof. Let $\gamma < 2$. If $\sigma_2(n) < \gamma n$, then there clearly is an l so that $\sigma(n) \not\equiv 0 (\operatorname{mod} 2^l)$ or n has fewer than l prime factors which occur in the factorization of n with an exponent 1. In other words, $n = R_1 R_2$, $(R_1, R_2) = 1$, where R_1 is square free and has fewer than l prime factors and all prime factors of R_2 occur with an exponent greater than 1. From this remark it follows by a simple computation that if $\gamma < 2$, there is an $l = l(\gamma)$ such that

$$N_\sigma(2, \gamma, x) < c_3 \frac{x(\log\log x)^{l-1}}{\log x}.$$

By the methods used in the proof of Theorem 1 it is easy to show that for every $\gamma > \frac{3}{2}$

$$N_\sigma(2, \gamma, x) > c_4 \frac{x}{\log x}.$$

We do not give the details of the proof.

If $\sigma_2(n) < \frac{3}{2}n$, then n and $\sigma(n)$ must be odd; hence n is a square and thus $N_\sigma(2, \frac{3}{2}, x) < x^{1/2}$. In fact, it would be easy to show that $N_\sigma(2, \frac{3}{2}, x) = o(x^{1/2})$ and $N_\sigma(2, \frac{3}{2}, x) > c_5 x^{1/2}/\log x$. It will not be easy to obtain an asymptotic formula for $N_\sigma(2, \frac{3}{2}, x)$. Similarly, we could investigate $N_\sigma(2, a, x)$ for $a < \frac{3}{2}$. We only make one final remark. It is easy to prove that if $n_1 < n_2 < \ldots$ is a sequence of integers for which $\sigma_2(n_i)/n_i \to 1$, then, for every $\varepsilon > 0$, $\sum_{n_i \leq x} 1 = o(x^\varepsilon)$.

Now we prove Theorem 1. First we prove the first inequality in (2). We need the following

LEMMA. *To every $\eta > 0$ there is a $c_\eta > 0$ such that the number of primes $p < x$ for which*

$$\frac{\varphi(p-1)}{p-1} < \frac{1-\eta}{2} \tag{8}$$

is greater than $c_3 x/\log x$.

A simple computation shows that (8) holds if (r odd prime)

$$\sum_{r|p-1} \frac{1}{r} < \eta. \tag{9}$$

Thus, to prove our lemma it will suffice to show that the number of primes $p < x$ satisfying (9) is greater than $c_\eta x/\log x$. To see this let $k = k(\eta)$ be sufficiently large and let $3 = q_1 < \ldots < q_k$ be the first k odd primes. Let $p_1 < \ldots < p_1 \leqq x$ be the set of primes $p < x$ satisfying $p \equiv -1 \left(\bmod \prod_{j=1}^{k} q_j\right)$. It follows from the prime number theorem for arithmetic progressions that

$$l = \bigl(1 + o(1)\bigr) \frac{x}{\log x} \prod_{j=1}^{k} (q_j - 1)^{-1}. \tag{10}$$

Now we prove

$$(11) \qquad \sum_{i=1}^{l} \sum_{r|p_i-1} \frac{1}{r} < \frac{1}{2}\eta_1 l.$$

If $r|p_i-1$, we must have $p_i \equiv -1 \pmod{\prod_{j=1}^{k} q_j}$ and $p_i \equiv 1 \pmod{r}$. By a theorem of Titchmarsh-Prachar ([3], p. 44, Theorem 4.1) the number of those primes $A(r, x)$ not exceeding x is less than

$$(12) \qquad c_6 \frac{x}{r\prod_{j=1}^{k}(q_j-1)} \log\left(\frac{x}{r\prod_{j=1}^{k} q_j}\right)^{-1}.$$

From (12) and (10) we obtain by a simple calculation (clearly $r|p_i-1$ implies $r > q_k$)

$$\sum_{i=1}^{l} \sum_{r|p_i-1} \frac{1}{r} = \sum_{q_k<r\leqslant x} \frac{A(r, x)}{r}$$
$$< c_6 \sum_{q_k<r\leqslant x} \frac{x}{r^2\prod_{j=1}^{k}(q_j-1)} \left(\log \frac{x}{r\prod_{j=1}^{k} q_j}\right)^{-1} < \frac{1}{2}\eta_1 l,$$

which proves (11). From (11) we immediately deduce that the number of primes $p_i < x$ which satisfy (9) is greater than $l/2$, which by (10) proves our lemma.

Let now $\alpha < \frac{1}{2}$ be given and choose $\eta = \eta(\alpha, t)$ to be sufficiently small. Let $p_1' < p_2' < \ldots$ be the primes satisfying (8) where $p_1' > c(\eta, t)$. By our lemma we have for $y > y(\eta, t)$

$$(13) \qquad \sum_{p_i'<y} 1 > \frac{1}{2} c_\eta \frac{y}{\log y}.$$

Denote by $u_1 < u_2 < \ldots$ the integers composed of at most $t+2$ primes p_i'. From (13) we infer by a simple computation using induction with respect to t that ($c_7 = c_7(\eta)$)

$$(14) \qquad \sum_{u_i<x} 1 > c_7 \frac{x(\log\log x)^{t+1}}{\log x}.$$

From (8) we obtain

$$(15) \qquad \varphi_2(u_i) > \frac{1}{2}(1-\eta)^t \varphi(u_i)$$

and from $p_1' > c(\eta, t)$ we have

$$(16) \qquad \varphi(u_i) > u_i\left(1 - \frac{1}{c(\eta, t)}\right)^{t+2}.$$

(15) and (16) imply if η is sufficiently small and $c(\eta, t)$ sufficiently large that

$$(17) \qquad \varphi_2(u_i) > a u_i.$$

(14) and (17) prove the first inequality in (2).

Now we prove the second one. Let $k = k(a)$ be sufficiently large and let $q_1, \ldots, q_k$ be the first k primes. If $\varphi_2(n) > an$, we evidently have

$$(18) \qquad \sum_{p \mid \varphi(n)} \frac{1}{p} < \frac{1}{a} \quad \text{hence} \quad \sum_{q_i \mid \varphi(n)} \frac{1}{q_i} < \frac{1}{a}.$$

Hence by (18) and from the well-known theorem of Mertens ($\sum_{i=1}^{k} 1/q_i = \log\log k + O(1)$) we have for $k = k(a)$

$$(19) \qquad \varphi(n) \not\equiv 0 \pmod{q_{j_i}}, \quad j_1 < \ldots < j_r \leqslant k, \quad \sum_{i=1}^{r} \frac{1}{q_{j_i}} > \frac{1}{2} \log\log k.$$

There are clearly fewer than 2^k choices for $j_1 < \ldots < j_r \leqslant k$. Thus our proof will be complete if we show that for every choice of $j_1 < \ldots < j_r \leqslant k$ satisfying $\sum_{i=1}^{r} 1/q_{j_i} > \frac{1}{2}\log\log k$ the number of integers $n \leqslant x$ satisfying

$$(20) \qquad \varphi(n) \not\equiv 0 \pmod{q_{j_i}}, \quad j_1 < \ldots < j_r \leqslant k,$$

is less than

$$\frac{x}{\log x} (\log x)^{\varepsilon/2}$$

if $k = k(\varepsilon, a)$ is sufficiently large.

It is easy to see that (20) implies that every prime factor p of n satisfies $p \not\equiv 1 \pmod{q_{j_i}}$, $j_1 < \ldots < j_r \leqslant k$. From the prime number theorem for arithmetic progressions and the sieve of Eratosthenes using (19) we easily obtain that the set of primes $s_1 < s_2 < \ldots$ for which $s \not\equiv 1 \pmod{q_{j_i}}$, $i = 1, \ldots, r$, satisfies

$$(21) \qquad \sum_{s_i \leqslant x} \frac{1}{s_i} = (1 + o(1)) \prod_{i=1}^{r} \left(1 - \frac{1}{q_{j_i}}\right) \log\log x$$

$$< \exp\left(-\sum_{i=1}^{r} \frac{1}{q_{j_i}}\right) \log\log x < \frac{\varepsilon}{4} \log\log x$$

if $k = k(\varepsilon)$ is sufficiently large.

If n satisfies (20), it must be composed entirely of the s_i's. Hence if $t_1 < t_2 < \ldots$ are the primes $\leqslant x$ which are not s_i's, we must have $n \not\equiv 0 (\bmod\ t_j)$. From (21) we have

$$\sum \frac{1}{t_j} > \left(1 - \frac{\varepsilon}{4}\right) \log\log x. \tag{22}$$

From (22) we deduce by Brun's method that the number of these $n \leqslant x$ is less than (if $x > x_0(\varepsilon)$)

$$c_8 x \prod_{t_j < x} \left(1 - \frac{1}{t_j}\right) < \frac{x}{\log x} (\log x)^{\varepsilon/2}$$

which completes the proof of (2).

To complete the proof of Theorem 1 we now have to prove (3). We will only outline the proof, since it is similar to the proof of the second part of (2). If $\varphi_3(n) > an$, we must have $\sum_{p | \varphi_2(n)} 1/p < 1/a$; hence, as in the previous proof, we must have (as in (19))

$$\begin{gathered} \varphi_2(n) \not\equiv 0 (\bmod q_{j_i}), \\ j_1 < \ldots < j_r \leqslant k, \quad \sum_{i=1}^{r} \frac{1}{q_{j_i}} > \frac{1}{2} \log\log k. \end{gathered} \tag{23}$$

Denote, as in the previous proof, by $t_1 < t_2 < \ldots$ the primes for which $t \equiv 1 (\bmod q_{j_i})$ for some j_i, $i = 1, \ldots, r$, and by $s_1 < s_2 < \ldots$ the set of primes for which

$$s \not\equiv 1 (\bmod t_j), \quad j = 1, 2, \ldots \tag{24}$$

(23) clearly implies that n is composed entirely of the s_i.

From (24) and (22) it follows by Brun's method that for $y > y_0(\varepsilon)$

$$\sum_{s_i < y} 1 < \frac{y}{(\log y)^2} (\log y)^{\varepsilon/2}. \tag{25}$$

We need the following

LEMMA. *Let $\{s_i\}$ be a sequence of primes satisfying* (25). *Then the number of integers not exceeding x of the form $\prod s_i^{a_i}$ is less than*

$$\frac{c_9 x}{(\log x)^2} (\log x)^{\varepsilon/2}.$$

We supress the details of the proof.

Since there are fewer than 2^k choices for $j_1 < \ldots < j_r \leqslant k$, our lemma immediately implies (3) and hence the proof of Theorem 1 is complete.

By the same method we can prove that

$$N_\varphi(4, a, x) < \frac{c_{10}x}{(\log x)^2}, \tag{26}$$

where c_{10} is an absolute constant independent of a.

(26) is probably very far from being the best possible.

REFERENCES

[1] Ju. V. Linnik, *The dispersion method in binary additive problems*, American Mathematical Society, Providence, Rhode Island 1963.

[2] A. Mąkowski and A. Schinzel, *On the functions* $\varphi(n)$ *and* $\sigma(n)$, Colloquium Mathematicum 13 (1964), p. 95-99.

[3] K. Prachar, *Primzahlverteilung*, Berlin-Göttingen-Heidelberg 1957.

Reçu par la Rédaction le 20. 1. 1966